地球科学大数据研究主题演化与影响力分析

安培浚[1,3*]　肖仙桃[1]　郭华东[2]　闫冬梅[2]　李佳蕾[1,3]
1 中国科学院西北生态环境资源研究院 兰州 730000
2 中国科学院空天信息创新研究院 北京 100094
3 中国科学院大学 经济与管理学院图书情报与档案管理系 北京 100049

摘要　大数据在地球科学各个学科中的应用越来越受到关注，数据驱动地球科学发现的案例不断出现，有关地球数据信息中心、地球大数据平台及相关学术会议数量逐渐增加，地球大数据正在科学研究上表现出巨大的潜力。科学家对地球大数据的科学方法和工具的需求很大，然而目前地球大数据的理论基础、储存管理和分析方法等仍处于发展之中，对地球大数据的研究和讨论有限。文章通过文献计量学的方法，对科学引文索引（SCI）和社会科学引文索引（SSCI）收录的地球大数据相关文献进行分析，从全球论文的产出数量、国家与机构领域研究影响力、研究主题分布、研究热点变迁和国际合作等多角度，分析揭示了地球大数据研究现状；最后，建议未来重点加强跨学科的地球大数据共享与融合，完善地球科学大数据深度挖掘理论和方法，实现对复杂地球系统的分析、建模与预测，支持和服务全球变化与可持续发展。

关键词　地球科学大数据，文献计量，研究主题分布，研究热点演化，研究影响力

DOI　10.16418/j.issn.1000-3045.20210807001

1998年时任美国副总统戈尔提出数字地球概念——将数字地球看成是对地球的三维多分辨率表示，并能放入大量地理数据。自此，地球大数据概念和相关研究开始走进大众视野。而在科学研究领域，地球大数据的研究及其应用其实早已开始。1954年，基于系列海洋数据分析了热比容偏差；1959年，基于卫星数据对大气进行分析；随着美国陆地（Landsat）系列卫星的发射，1976年起地球科学与大数据相关研究论文明显增多，大多为基于卫星数据开展分析，主要应用于地质和矿产资源勘探、作物面积和土地利用

* 通信作者
项目资助：中国科学院战略性先导科技专项（A类）（XDA19000000），中国科学院政策研究课题（ZYS-2021-07）
修改稿收到日期：2021年8月11日

分类估算、大气成分和云层状况研究、冰冻预报、卫星和遥感数据处理技术等。

随着计算机互联网、卫星定位系统、空间信息技术、高分辨率卫星影像、大数据处理与存储技术、科学计算及可视化、虚拟现实技术等现代信息科技的迅猛发展，地球大数据的深入应用将对全球和各国的科技、经济和社会生活产生巨大影响。因此，地球大数据相关理论与技术方法研究成为当前研究热点领域[1-3]。本文基于文献计量方法，对地球大数据研究主题演化与影响力进行分析，以期为相关领域研究提供参考。

1 数据源与研究方法

利用“地球科学”和“大数据”主题组合检索策略构建检索词和检索式（广义的），采集文献覆盖面广与影响力大的科学引文索引（SCI）和社会科学引文索引（SSCI）数据库的所有文献数据（截至2021年7月20日）。经过对机构和关键词等信息的清洗，并经专家识别排除不相关文献后，利用多种可视化分析工具，定量分析地球大数据研究的现状与研究主题变化情况。

2 地球大数据研究总体科研产出分析

2.1 全球地球大数据论文增长趋势

地球大数据研究年发文量变化可以反映出其研究变迁与研究热度。图1展示了全球地球大数据相关研究年度发文量的变化趋势。数据显示，1990年以前，地球大数据相关研究论文还非常少；1991—2000年，论文数量开始保持小幅增长；2000年以来，论文增长进入快车道；2017年之后更呈现暴发式增长态势。

地球大数据研究论文占全球论文比重可以更直观地反映该领域受全球关注程度。图2统计了1954—2020年地球大数据相关论文占SCI/SSCI年出版论文的比重，1990年以后，地球大数据相关研究论文占全球论文比重呈现逐步增长的态势；特别是2012年之后，

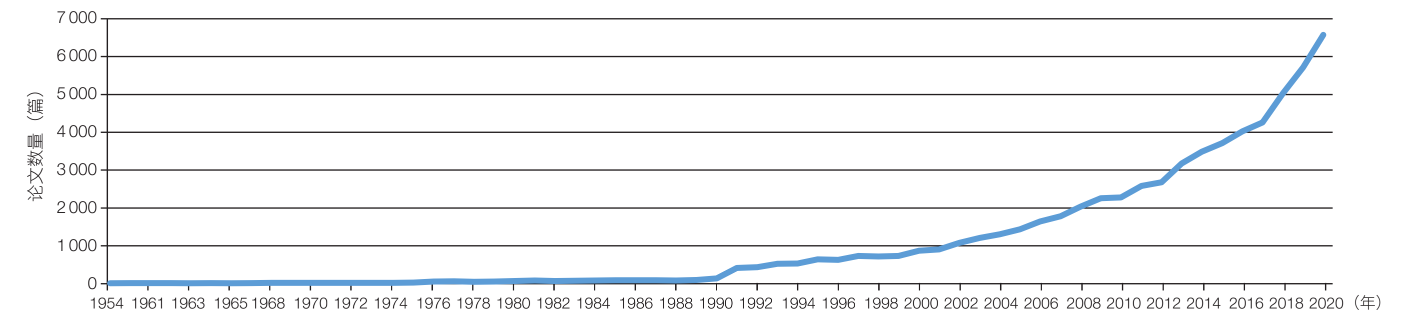


图1 1954—2020年地球大数据相关论文数量增长趋势图

Figure 1 Trends in number of papers related to Big Earth Data from 1954 to 2020

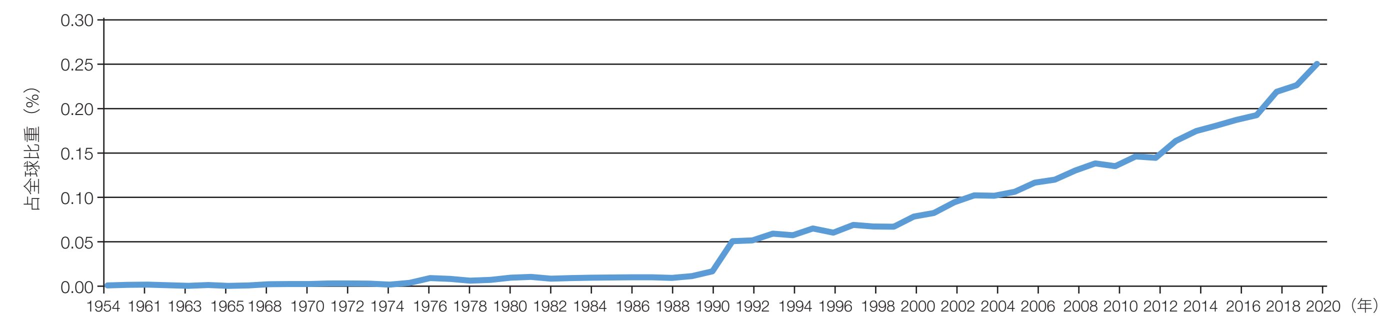


图2 1954—2020年地球大数据相关论文数量占全球论文比重变化趋势

Figure 2 Changing trend of percentage of number of papers related to Big Earth Data from 1954 to 2020

增长更加迅猛；2020年，地球大数据相关研究论文占全球比重已达到0.25%。

2.2 全球国家与机构的科研影响力分析

地球大数据相关研究论文主要分布在全球近200个国家/地区。表1基于论文总被引频次列出论文影响力最高的15个国家/地区及其论文数、篇均被引频次、论文数全球占比和近10年（2011—2021年）美国科技信息研究所基本科学指标数据库（ESI）前1%高被引论文全球占比。美国涉及地球科学的大数据研究（广义领域）起步较早，在论文数、论文总体影响力、篇均被引频次、高被引论文等指标上都表现较好。首次明确提出地球大数据与大数据科学的概念（具体领域）是中国的科学家。2015年6月17—18日，在北京举行的“空间地球大数据”科学与技术前沿论坛上，郭华东报告中首次探讨了空间地球大数据理念、科学内涵、与数字地球关系，以及对推动地球系统科学及全球变化研究发展的潜力；2016年，郭华东及其团队在《科学通报》发表的《空间地球观测数据成为地球科学新引擎》一文正式使用地球大数据概念；2017年，郭华东牵头创办了全球地球科学领域的第一个大数据刊物 *Big Earth Data*，并在创刊号上发文，提出地球大数据科学的概念。目前，中国已成为地球科学大数据的研究大国，年相关论文产出超过美国，在国际地球大数据科学领域的研究已处于引领地位。

图3可视化展示了地球大数据相关研究论文产出最多的15个国家/地区的年度论文数量变化，主要国家/地区相关论文产出均呈增长态势。2000年以前，美国发表的相关论文占全球至少一半以上；2010年以后，中国相关论文增长很快，德国、澳大利亚、英国

表1 1954—2020年地球大数据相关研究论文影响力最高的15个国家/地区

Table 1 Top 15 countries/regions with high influence of papers related to Big Earth Data from 1954 to 2020

国家/地区	论文数（篇）	总被引频次	篇均被引频次	论文数占全球比（%）	ESI前1% 论文数占全球比（%）
美国	22 674	1 041 038	45.91	32.83	50.25
中国	14 123	281 079	19.90	20.45	30.35
英国	5 437	221 049	40.66	7.87	14.36
德国	5 196	175 464	33.77	7.52	13.48
印度	3 974	68 416	17.22	5.75	3.90
法国	3 873	154 874	39.99	5.61	10.20
意大利	3 752	130 355	34.74	5.43	10.33
加拿大	3 720	136 017	36.56	5.39	9.82
澳大利亚	3 162	114 884	36.33	4.58	11.34
西班牙	2 353	78 763	33.47	3.41	7.18
日本	2 334	67 499	28.92	3.38	4.66
荷兰	2 139	95 402	44.60	3.10	8.19
巴西	1 720	45 618	26.52	2.49	3.15
俄罗斯	1 565	26 211	16.75	2.27	1.39
韩国	1 367	29 667	21.70	1.98	3.02

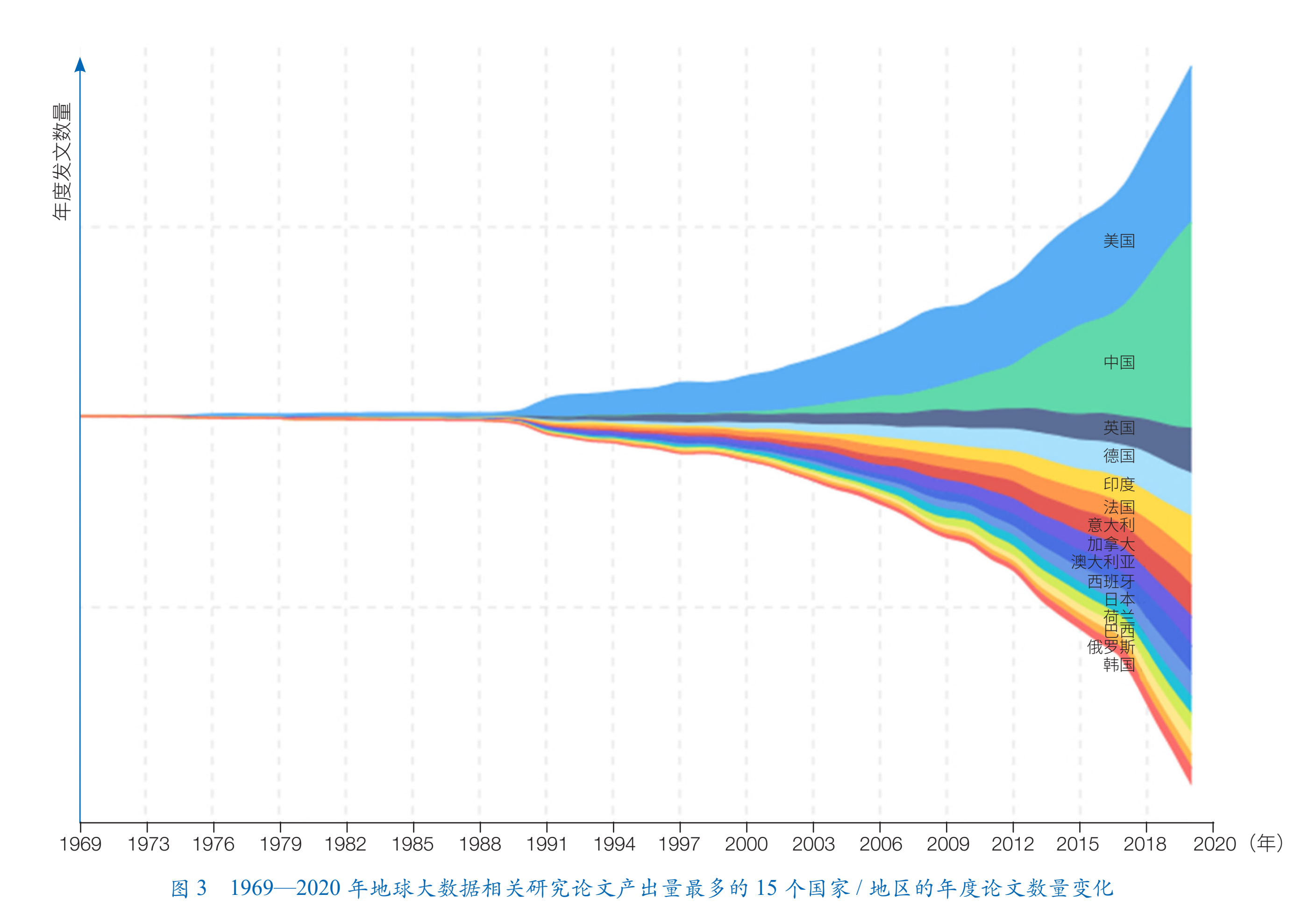


图 3　1969—2020 年地球大数据相关研究论文产出量最多的 15 个国家 / 地区的年度论文数量变化

Figure 3　Annual paper number change of top 15 countries/regions with papers related to Big Earth Data from 1969 to 2020

等国相关论文也有明显增长；自 2018 年起，中国年相关论文年产出已超过美国。

表 2 分别统计了全球地球大数据领域发文最多和论文总被引频次最高的 25 个机构。发文最多的 25 个机构分布在美国、中国、法国、俄罗斯、意大利、印度等国，其中美国 13 个，中国 8 个，法国、俄罗斯、意大利、印度各 1 个；论文总被引频次最高的 25 个机构分布在美国、中国、法国、意大利、加拿大等国，其中美国 19 个，中国 3 个，法国、意大利、加拿大各 1 个。

中国在地球大数据领域发文从论文数量和论文总被引频次来看，中国科学院、北京师范大学和武汉大学是国内 3 个最主要的研究机构。北京师范大学开展地球大数据研究的主要单元为中国科学院遥感与数字地球研究所和北京师范大学联合建设的遥感科学国家重点实验室；同时，北京师范大学地理科学学部设置了地理数据与应用分析中心，开展地理空间数据建设、计算平台搭建，以及利用多源数据开展地理学过程分析，中心下设数据存储与计算实验室、地理数据服务研究所和地理大数据研究所。武汉大学遥感信息工程学院和测绘学院是该校开展地球大数据研究的主要研究单元。

3 地球大数据研究领域分布

3.1 论文所属期刊情况

基于美国科学信息研究所《期刊引用报告》（*JCR*）期刊分类对地球大数据相关研究论文所属期刊学科领域进行分析发现，地球大数据相关研究论文所

表 2　全球地球大数据领域发文最多与论文总被引频次最高的 25 个机构

Table 2　Top 25 institutions in paper numbers or high citations related to Big Earth Data

序号	发文最多			论文总被引频次最高		
	机构	所在国家	论文数量（篇）	机构	所在国家	论文总被引频次（次）
1	中国科学院	中国	5 157	美国国家航空航天局	美国	176 749
2	美国国家航空航天局	美国	2 435	中国科学院	中国	118 362
3	法国国家科学研究中心	法国	1 903	马里兰大学	美国	113 143
4	马里兰大学	美国	1 655	美国海洋与大气管理局	美国	78 984
5	美国海洋与大气管理局	美国	1 253	法国国家科学研究中心	法国	66 464
6	北京师范大学	中国	1 214	美国地质调查局	美国	52 569
7	武汉大学	中国	1 111	美国国家大气科学研究中心	美国	47 138
8	美国地质调查局	美国	998	科罗拉多大学	美国	46 612
9	俄罗斯科学院	俄罗斯	885	威斯康星大学	美国	43 940
10	意大利国家研究委员会	意大利	884	波士顿大学	美国	43 623
11	加州理工学院	美国	791	北京师范大学	中国	34 862
12	威斯康星大学	美国	763	加州理工学院	美国	34 091
13	科罗拉多大学	美国	677	加州大学圣芭芭拉分校	美国	32 158
14	印度理工学院	印度	664	科学应用国际公司	美国	27 785
15	美国农业部农业研究局	美国	568	美国农业部农业研究局	美国	27 103
16	北京大学	中国	555	亚利桑那州立大学	美国	26 564
17	南京大学	中国	503	意大利国家研究委员会	意大利	26 368
18	美国农业部林业局	美国	469	哥伦比亚大学	美国	26 215
19	南京信息工程学院	中国	441	俄勒冈州立大学	美国	25 651
20	中国气象局	中国	454	华盛顿大学	美国	25 143
21	华盛顿大学	美国	435	美国农业部林业局	美国	25 119
22	清华大学	中国	433	武汉大学	中国	24 058
23	得克萨斯农工大学	美国	425	加州大学伯克利分校	美国	19 934
24	俄勒冈州立大学	美国	421	不列颠哥伦比亚大学	加拿大	19 381
25	波士顿大学	美国	404	俄亥俄州立大学	美国	18 713

属期刊分布在 200 多个主题类目中，大部分为自然科学，也有少量为社会科学和人文科学，图 4 为分布最多的 40 个主题类目。遥感、环境科学、影像科学与摄影技术、地球科学多学科、气象与大气科学为分布最多的 5 个主题领域。

根据统计，地球大数据相关论文分布在 SCI 收录的 4 000 余种期刊、丛书、图书中。其中，发文最多的 20 种期刊中发表的相关论文占全部相关论文一半以上（图 5）。*Remote Sensing*、*International Journal of Remote Sensing*、*Remote Sensing of Environment*、*Journal of Geophysical Research emdash Atmospheres*、*IEEE Transactions on Geoscience and Remote Sensing* 是发表相

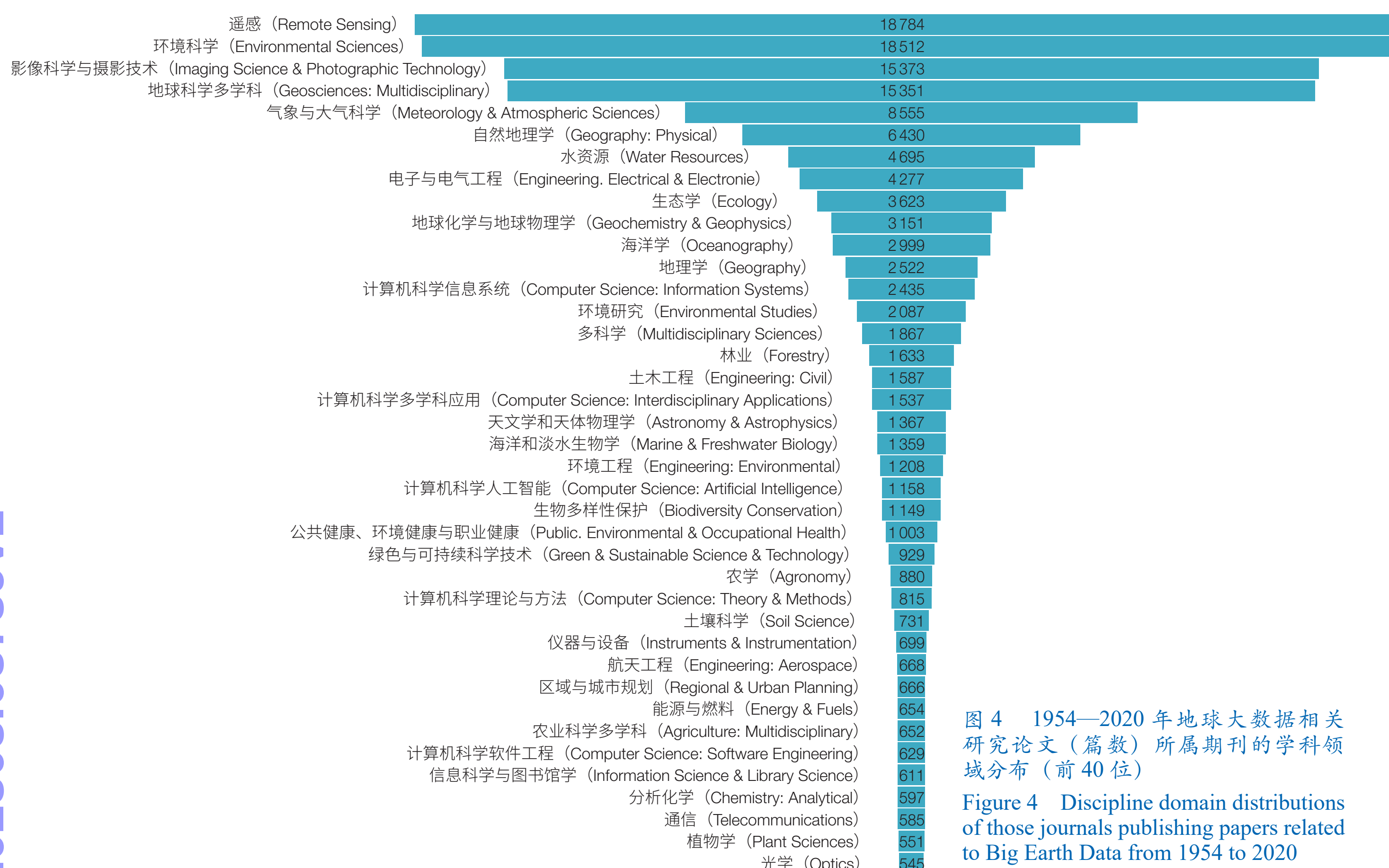


图 4　1954—2020 年地球大数据相关研究论文（篇数）所属期刊的学科领域分布（前 40 位）

Figure 4　Discipline domain distributions of those journals publishing papers related to Big Earth Data from 1954 to 2020

关论文最多的 5 个期刊。

3.2 论文所属研究主题情况

InCites①开发了一种基于论文引用的分类算法。随着已发表论文的相互引用，这些引用关系的强度将相关文献汇聚到一起，形成离散的相关文献集群。这些集群构成了引文主题（Citation Topics）的核心，独立于单篇文献的主题和内容，却代表着作者相互积极引用对方论文的领域。Citation Topics 构建一个宏观、中观和微观主题的 3 级层次结构，包含 10 个广义集群、326 个中观集群和 2 444 个微观集群的层次结构。根据聚类内容对宏观和中观类别进行了标注，由于微观 Citation Topics数量较多，使用算法工具根据聚类文献中最重要的关键词给每个主题加标注，从而有助于对研究主题进行更精细的分析。基于 InCites 的 Citation Topics 分类，对地球大数据相关研究论文的主题分布进行分析。

（1）基于 Citation Topics的宏观领域分布。地球大数据相关研究论文主要分布在电气工程、电子学与计算机科学，地球科学，农业、环境与生态三大领域。这表明信息技术是地球大数据研究最主要的研究基础；而地球科学，农业、环境和生态领域是地球大数据的主要应用领域（图 6）。

① InCites 是汤森路透科技集团在汇集和分析 Web of Science（SCI/SSCI）权威引文数据的基础上建立起来的科研评价工具，综合各种计量指标和 30 年来各学科各年度的国际标杆数据。通过 InCites，用户能够实时跟踪机构的研究产出和影响力；将某机构的研究绩效与其他机构、全球和学科领域的平均水平进行对比；发掘机构内具有学术影响力和发展潜力的研究人员，并监测机构的科研合作活动，以寻求潜在的科研合作机会。InCites 能够帮助政府和学术研究机构中的决策者、科研管理人员分析本机构的学术表现和影响力，并针对全球同行的研究成果进行比较。

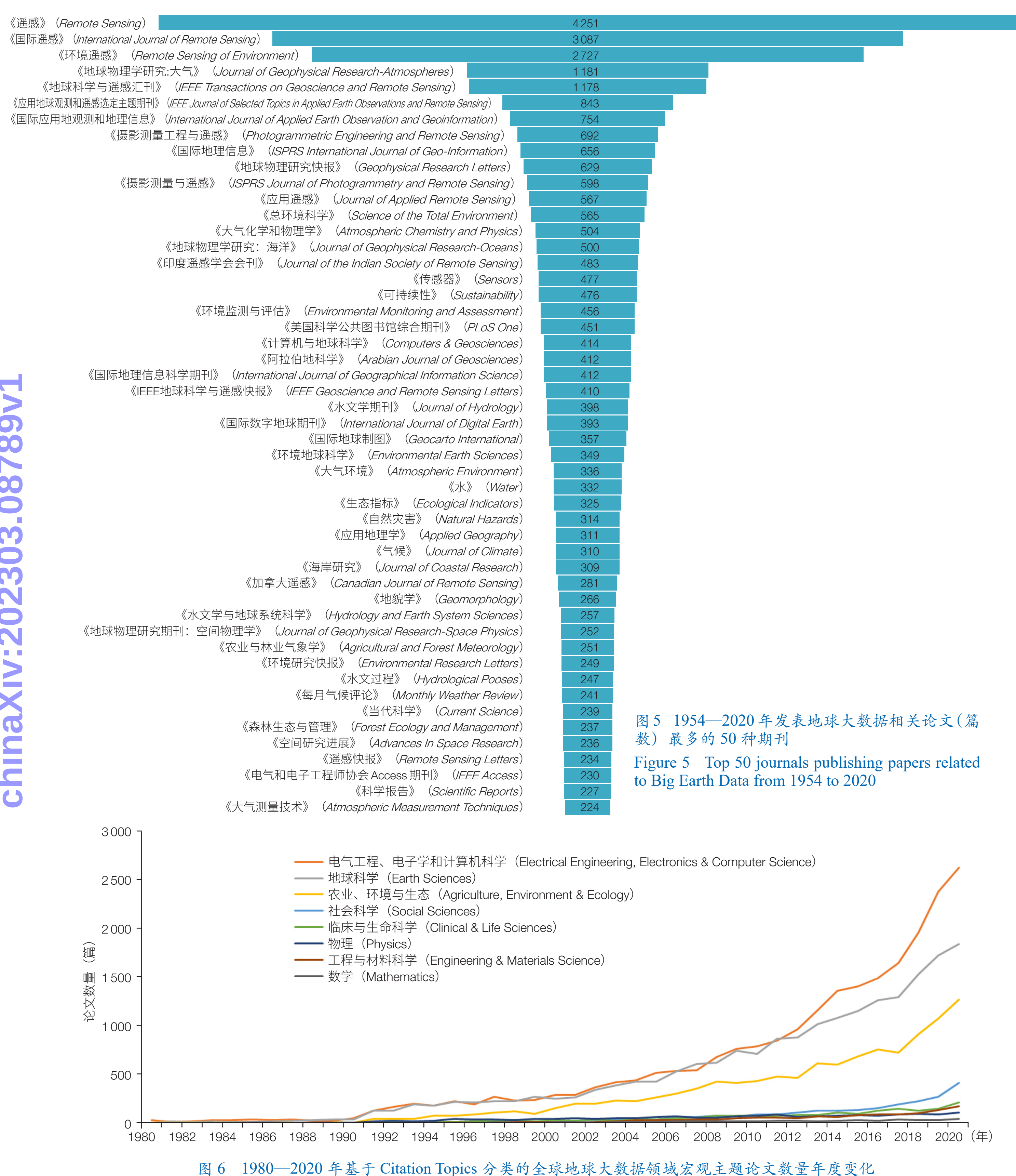


图5　1954—2020年发表地球大数据相关论文（篇数）最多的50种期刊

Figure 5　Top 50 journals publishing papers related to Big Earth Data from 1954 to 2020

图6　1980—2020年基于Citation Topics分类的全球地球大数据领域宏观主题论文数量年度变化

Figure 6　Annual paper number change of macro themes related to Big Earth Data based on Citation Topics classification from 1980 to 2020

（2）基于 Citation Topics 分类的地球大数据相关研究论文中观领域分布。研究主题主要集中在遥感，海洋学、气象与大气科学，林业等领域（图7）。

（3）基于 Citation Topics 分类的地球大数据相关研究论文微观领域分布。研究主题主要集中在归一化植被指数、蒸散发、气溶胶、海洋颜色、自发地理信息、冰川、云、生态系统服务、厄尔尼诺与南方涛动、热带气旋等（图8）。

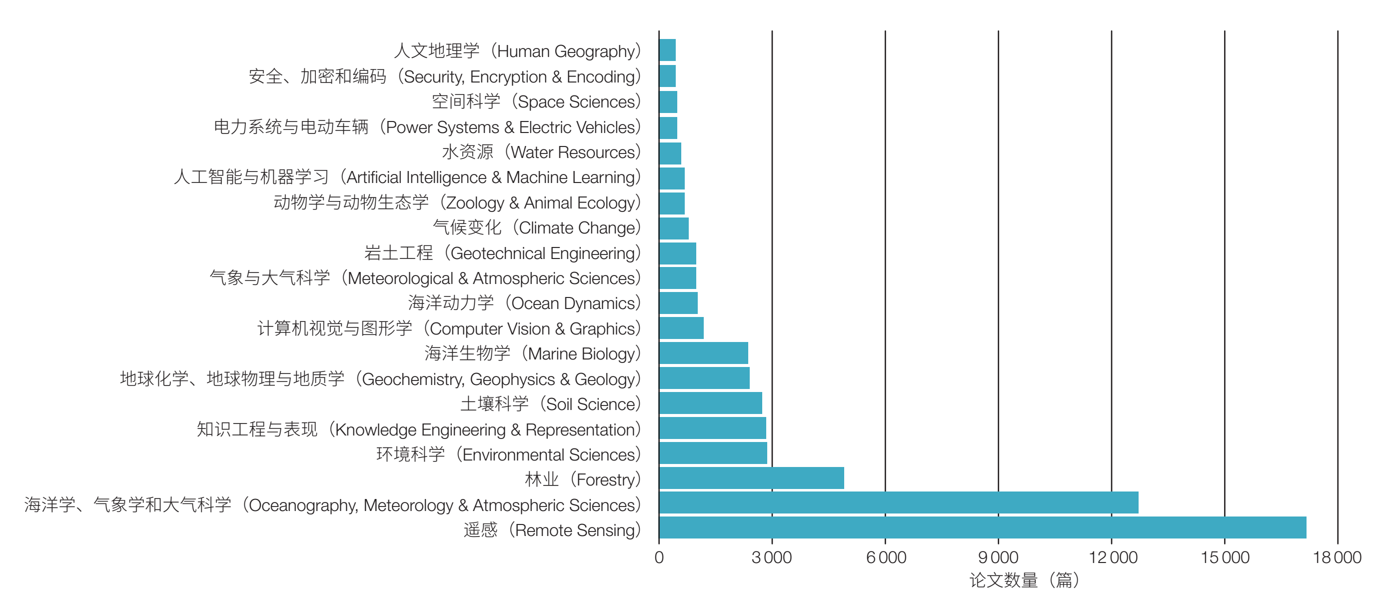


图7　1980—2020 年基于 Citation Topics 分类的全球地球大数据相关研究论文中观主题领域分布（前 20 位）

Figure 7　Top 20 meso themes related to Big Earth Data based on Citation Topics classification from 1980 to 2020

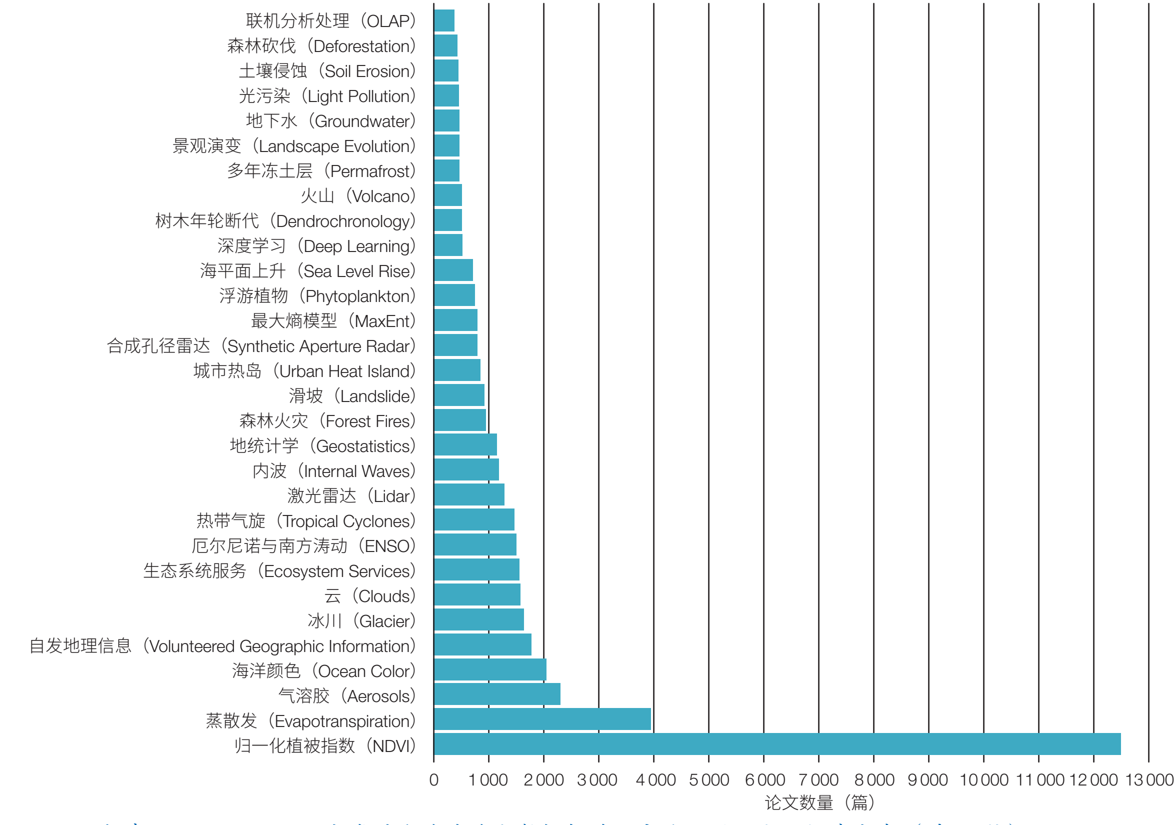


图8　1980—2020 年基于 Citation Topics 分类的全球地球大数据相关研究论文微观主题领域分布（前 30 位）

Figure 8　Top 30 micro themes related to Big Earth Data based on Citation Topics classification from 1980 to 2020

3.3 专业期刊关键词分布情况

SCI 数据库目前仅收录 2 种地球大数据领域专业期刊——*Earth System Science Data* 和 *Geoscience Data Journal*。*Earth System Science Data* 创刊于 2009 年，2012 年开始被 SCI 收录，目前为 JCR 地球科学“多学科”领域 1 区期刊，“气象与大气科学”领域 1 区期刊；*Geoscience Data Journal* 创刊于 2014 年，2015 年被 SCI 收录，目前该刊为 JCR 地球科学“多学科”领域 3 区期刊，“气象与大气科学”领域 4 区期刊（2020 版 JCR）。截至 2021 年 8 月，SCI 数据库共收录了以上 2 刊的 893 篇论文。

对 *Earth System Science Data* 和 *Geoscience Data Journal* 发表的 SCI 论文的主题分布进行分析，基于其所发表论文的“作者关键词”[②]，绘制了关键词共现网络。如图 9 显示，两刊发表论文研究主题涉及气候与天气、气候变化、大气温度、降雨、观测、重构、水文模型、海面温度等广泛主题，热点关键词主要有数据救援/气候数据救援（climate data rescue）、公民

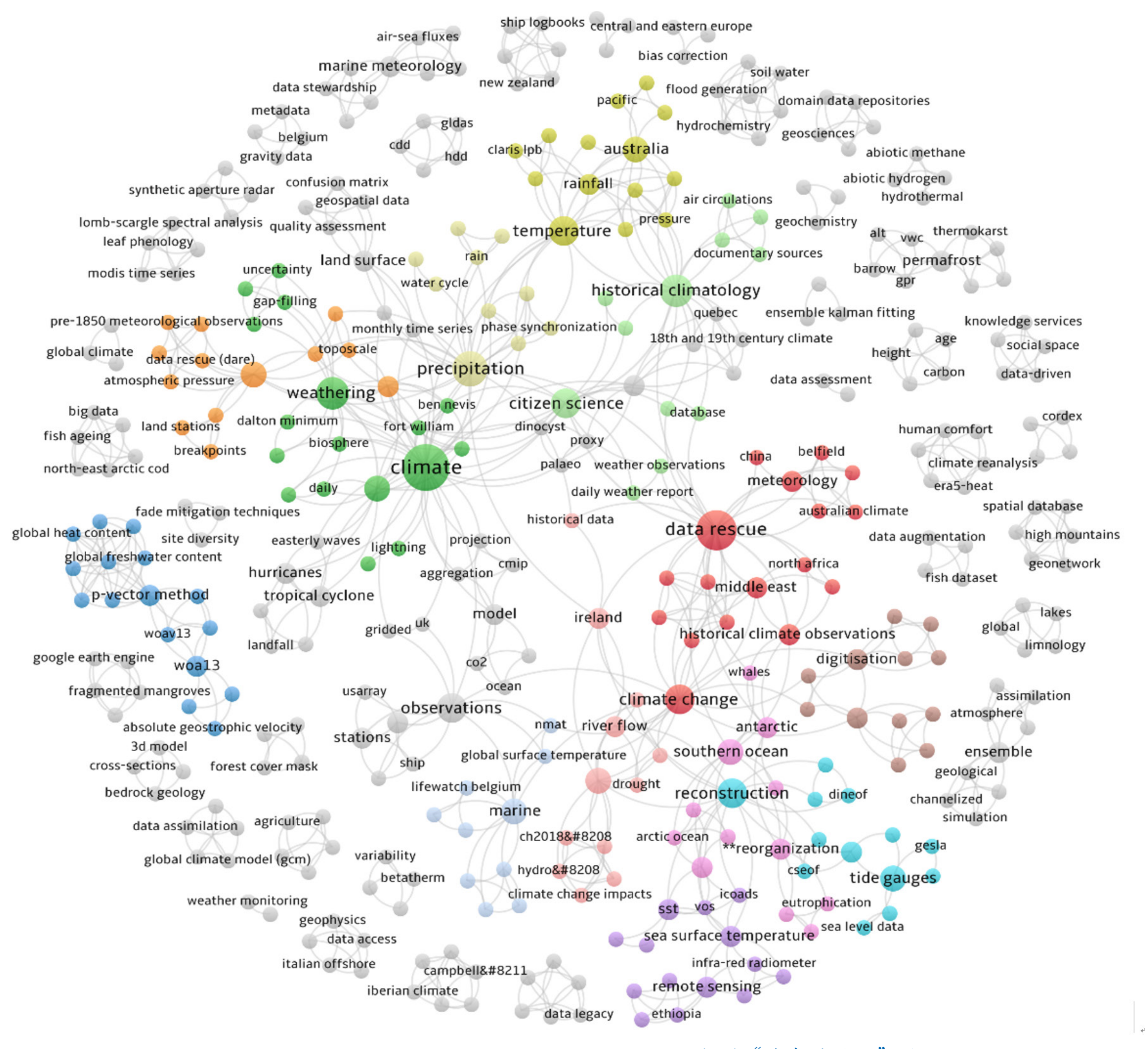


图 9　*Earth System Science Data* 和 *Geoscience Data Journal* 期刊“作者关键词”共现网络

Figure 9　Author keyword co-occurrence network of *Earth System Science Data* and *Geoscience Data Journal*

② 数据库中的关键词一部分是作者文章给出的关键词，称为“作者关键词”；另一部分是数据库商根据参考文献和论文内容补充的关键词，称为“标引关键词”。为更加准确起见，本文只分析了文章的作者关键词。

科学（citizen science）等。

4 地球大数据研究热点演化

提取不同时期相关论文最多的20个主题，并用热力色阶表示各主题在不同时期相关论文多少，颜色越接近红色表明文献量越大，颜色越接近蓝色表明文献越少（表3）。从表3可以看到，不同时期研究热点有较多不同。2000年以前，更关注归一化植被指数、云、厄尔尼诺与南方涛动、蒸散发、热带气旋、内波、自发地理信息、海洋的颜色、冰川、地统计学。

表3　不同时期主要热点主题包含的地球大数据相关研究论文数量对比（单位：篇）

Table 3　Comparison of paper numbers including main hot themes related to Big Earth Data in different periods

微观主题名称	2000年以前	2000—2009年	2010—2015年	2016—2020年
归一化植被指数（NDVI）	1 222	2 469	3 480	5 333
蒸散发（Evapotranspiration）	255	765	1 176	1 764
气溶胶（Aerosols）	100	553	750	929
海洋颜色（Ocean Color）	139	424	605	899
生态系统服务（Ecosystem Services）	51	319	438	772
冰川（Glacier）	127	289	495	754
自发地理信息（Volunteered Geographic Information）	141	345	580	742
激光雷达（Lidar）	38	182	361	720
城市热岛（Urban Heat Island）	33	97	174	565
地学统计（Geostatistics）	119	212	302	539
深度学习（Deep Learning）	0	8	24	513
热带气旋（Tropical Cyclones）	212	390	396	485
滑坡（Landslide）	19	168	278	451
森林火灾（Forest Fires）	43	236	253	437
云（Clouds）	341	428	397	425
厄尔尼诺与南方涛动（ENSO）	278	471	374	400
内波（Internal Waves）	157	338	323	394
合成孔径雷达（Synthetic Aperture Radar）	73	166	197	373
海平面上升（Sea Level Rise）	33	133	206	350
光污染（Light Pollution）	13	39	76	349
最大熵模型（MaxEnt）	57	220	223	308
浮游植物（Phytoplankton）	68	188	229	278
景观演变（Landscape Evolution）	52	110	137	196
联机分析处理（OLAP）	69	128	82	116
电离层（Ionosphere）	68	128	85	85
磁层（Magnetosphere）	88	156	90	58
定性空间推理（Qualitative Spatial Reasoning）	60	104	36	29

2016 年以来，更关注归一化植被指数、蒸散发、气溶胶、海洋颜色、生态系统服务、冰川、自发地理信息、激光雷达、城市热岛、地学统计；此外，深度学习、滑坡、森林火灾、光污染、城市热岛效应等主题的相关文献数量增长较快。

图 11—14 基于“作者关键词”可视化展示了不同时期相关研究论文的 300 个高频关键词聚类网络，从而可以一窥不同时期的主要热点研究主题。每幅图中用同一颜色代表同一个聚类。节点大小和节点字体大小表示该节点论文数量多少。可以看到不同时期热点

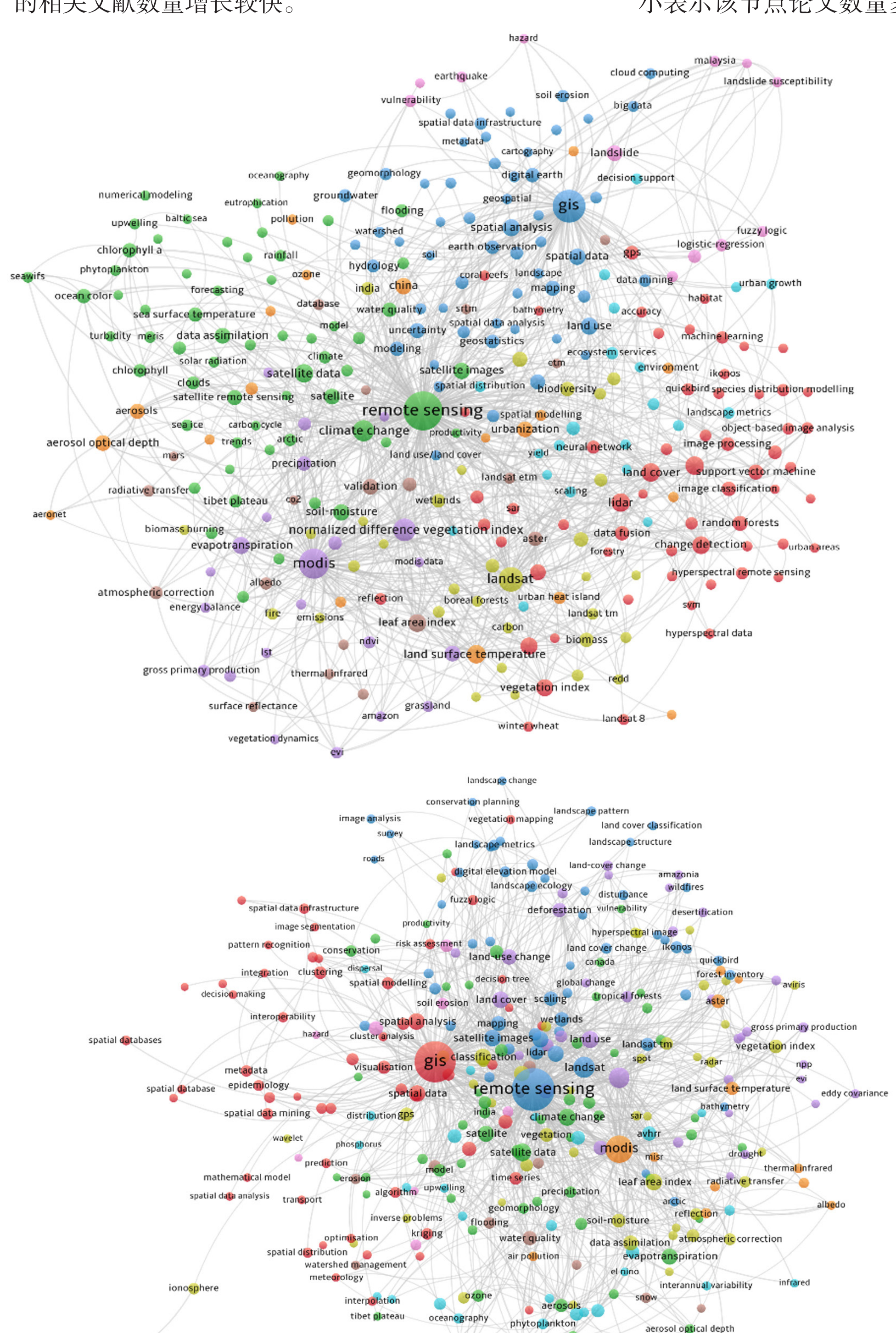


图 11　2000 年以前地球大数据相关研究论文高频关键词（前 300 个）共现网络

Figure 11　High-frequency keyword co-occurrence network of papers related to Big Earth Data before 2000

图 12　2000—2009 年地球大数据相关研究论文高频关键词（前 300 个）共现网络

Figure 12　High-frequency keyword co-occurrence network of papers related to Big Earth Data from 2000 to 2009

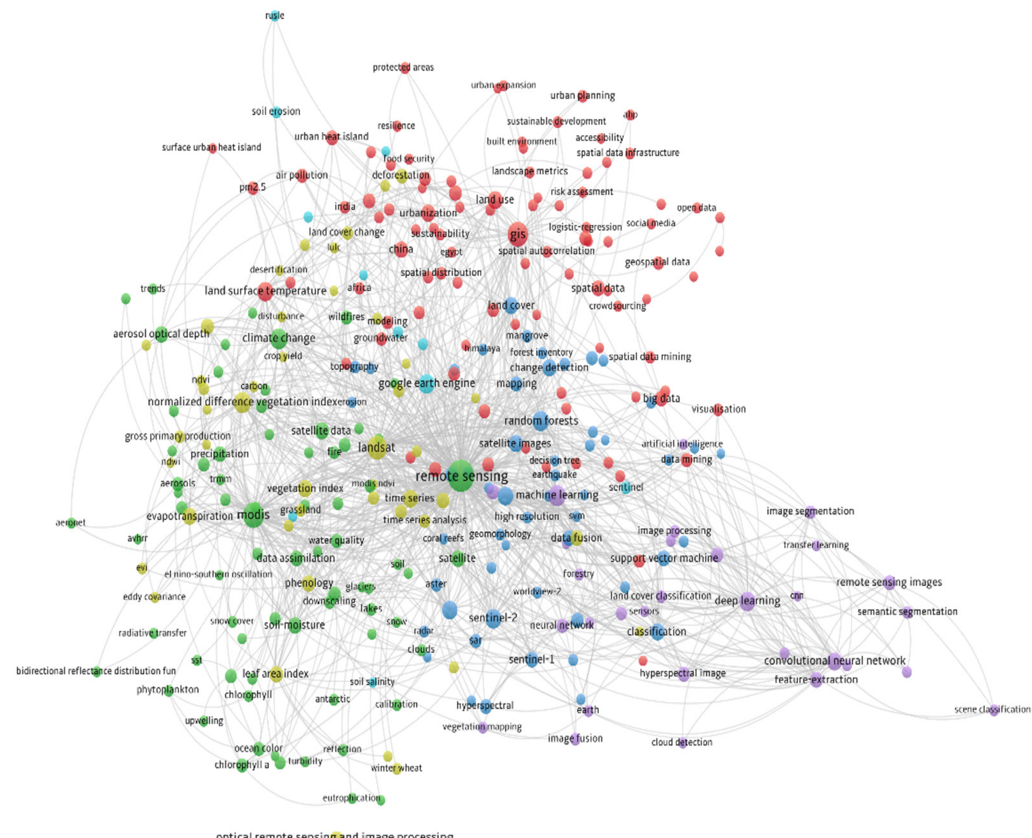


图 13　2010—2015 年地球大数据相关研究论文高频关键词（前 300 个）共现网络

Figure 13　High-frequency keyword co-occurrence network of papers related to Big Earth Data from 2010 to 2015

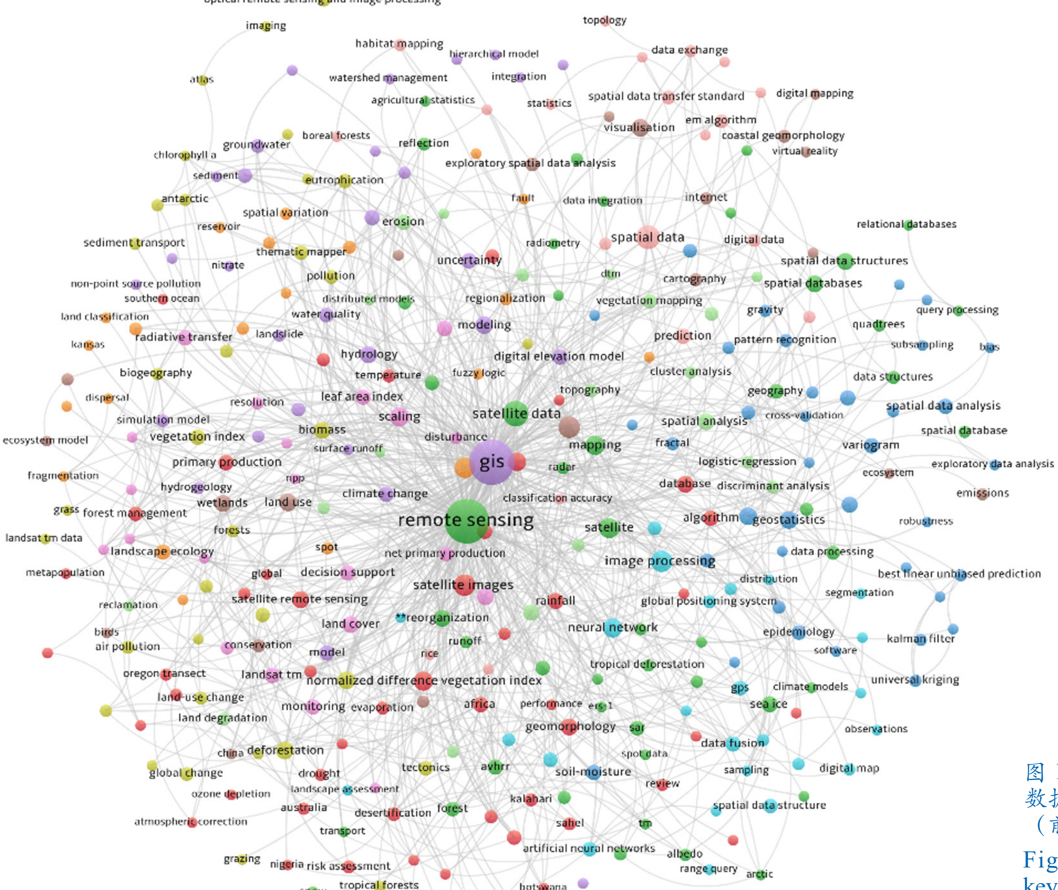


图 14　2016—2020 年地球大数据相关研究论文高频关键词（前 300 个）共现网络

Figure 14　High-frequency keyword co-occurrence network of papers related to Big Earth Data from 2016 to 2020

研究主题发生了一些变化。

5 地球大数据研究的国际合作情况

5.1 主要国家间的国际合作

图 15 可视化展示了地球大数据研究相关论文最多的 50 个国家/地区间合作密度，其中美国、中国、加拿大、英国、德国等国家/地区的国际合作强度最大。

表 4 统计了地球大数据相关研究论文发文量最多的 25 个国家/地区的国际合作论文占本国相关论文比重。数据显示，大部分国家/地区的国际合作论文所占比重均超过一半以上，国际合作论文比重占比较低的有印度、土耳其、中国。

5.2 中国主要国际合作伙伴

图 16 显示了在地球大数据领域，与中国开展相关研究论文国际合作最多的 25 个国家，其中连线粗细表示与中国合作论文数量多少。中国国际合作论文最多的国家是美国、英国、加拿大、澳大利亚、德国、法国、日本、荷兰、意大利、韩国、比利时。中国发表的地球大数据相关研究论文

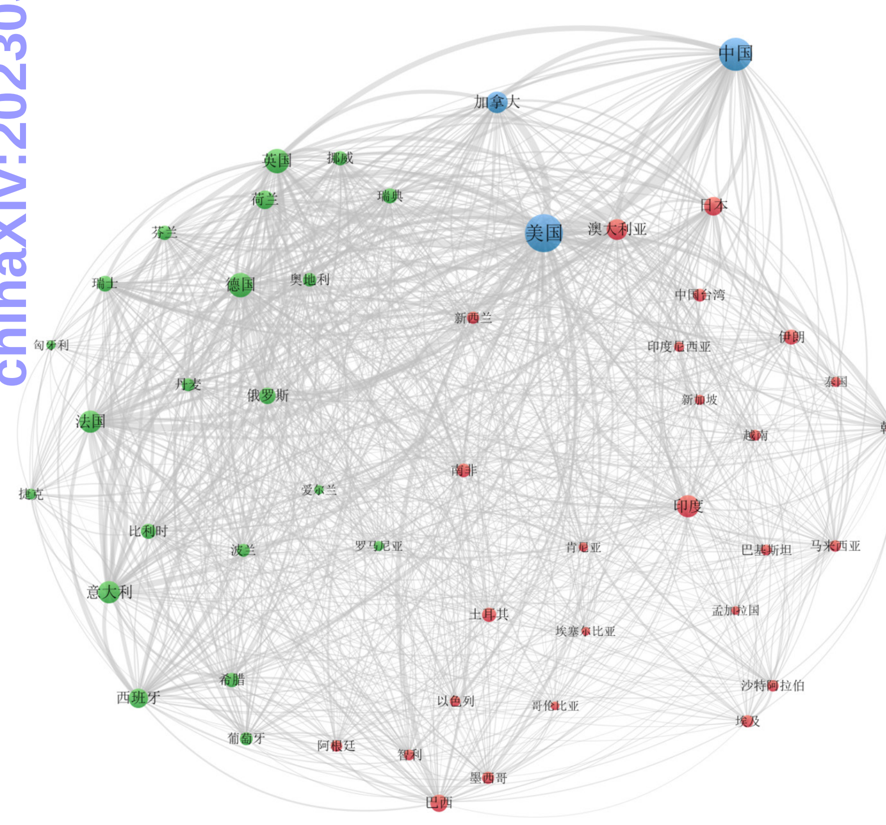


图 15　地球大数据相关研究论文产出国家 / 地区（前 50 位）间合作网络

Figure 15　Cooperation network of top 50 countries/regions in number of papers related to Big Earth Data

表4　地球大数据相关研究论文产出量最多国家/地区（前25位）的国际合作论文占比情况

Table 4　Percentage of international cooperation papers of top 25 countries/regions in number of papers related to Big Earth Data

排名	国家	国际合作论文占本国论文比重（%）
1	美国	42.66
2	中国	39.89
3	德国	63.37
4	英国	65.04
5	印度	23.71
6	意大利	55.34
7	法国	72.44
8	加拿大	55.08
9	澳大利亚	61.16
10	西班牙	61.49
11	荷兰	77.84
12	日本	55.50
13	巴西	50.15
14	韩国	46.04
15	伊朗	52.46
16	俄罗斯	40.39
17	瑞士	78.24
18	比利时	78.91
19	瑞典	69.61
20	土耳其	24.72
21	挪威	78.08
22	奥地利	76.39
23	芬兰	62.30
24	希腊	51.45
25	南非	60.31

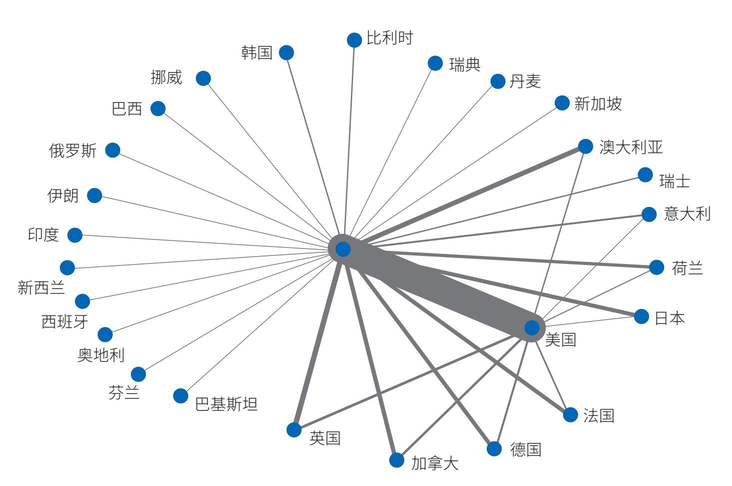


图 16　1954—2021 年地球大数据领域与中国开展相关研究论文国际合作最多的国家（前 25 位）

Figure 16　Top 25 countries in number of collaborative papers with China related to Big Earth Data from 1954 to 2021

中，近 10% 的论文为 3 国及以上国家合作完成。

表 5 统计了地球大数据相关研究中，与中国合作论文最多的 25 个国外机构。统计显示，中国的国际合作机构主要分布在美国，美国马里兰大学、法国国家科学研究中心、美国国家航空航天局是中国地球大数据领域最主要的 3 个海外合作机构。

6 结论与展望

近年来互联网、云计算和人工智能等相关技术发展，使得地球大数据研究热度升高，发文量增长迅速，影响力呈上升态势。美国和中国在地球大数据研究领域处于领先地位，且目前中国的年发文量已经超过美国，但高被引论文数量低于美国，论文质量发展提升空间较大。从全球地球大数据发文最多和论文总被引频次来看，研究热度较高的机构不一定是影响力较大的机构[4]。受我国地球大数据工程项目牵引，我国近 3 年的地球大数据发文量和影响力均得到快速提升，已经跃居全球第一，并且由我国科学家发起的基于地球大数据的“数字丝路”国际科学计划带动，与“一带一路”多个国家和国际组织合作，参与国家科研机构在地球大数据的科研

表 5　中国合作论文最多的25个国外机构

Table 5　Top 25 foreign institutions in number of collaborative papers with China related to Big Earth Data

排名	机构	所属国	合作论文数（篇）
1	马里兰大学	美国	352
2	法国国家科学研究中心	法国	184
3	美国国家航空航天局	美国	135
4	俄克拉荷马大学	美国	122
5	威斯康星大学	美国	99
6	密歇根州立大学	美国	94
7	乔治梅森大学	美国	92
8	美国国家海洋和大气管理局	美国	91
9	加州大学伯克利分校	美国	84
10	特文特大学	荷兰	77
11	多伦多大学	加拿大	75
12	美国农业部农业研究局	美国	68
13	东京大学	日本	63
14	波士顿大学	美国	61
15	美国地质调查局	美国	57
16	俄亥俄州立大学	美国	56
17	美国国家大气科学研究中心	美国	54
18	加州理工学院	美国	52
19	加州大学洛杉矶分校	美国	51
20	得克萨斯农工大学	美国	51
21	伊利诺伊大学	美国	48
22	昆士兰大学	澳大利亚	46
23	亚利桑那州立大学	美国	45
24	俄罗斯科学院	俄罗斯	44
25	加州大学圣芭芭拉分校	美国	43

产出上也得到明显的发展。地球大数据发文所属期刊来看，中国期刊的国际化与国际影响力仍有待提高，国际合作以华人居多，需要进一步深入拓展。地球大数据研究领域主要分布在遥感科学、环境科

学、大气科学等，其研究热点也主要聚焦在地球科学的研究热点主题上，以应用为主，理论方法主要涉及与信息技术相关的人工智能等。

尽管国内外地球科学大数据都已经形成了相当的规模，取得了可观的应用成果，但从发表的地球大数据相关研究论文内容来看，仍然存在诸多问题亟待解决，如：缺乏统一的数据描述标准[5]、规范的数据共享机制，以及数据融合与深度挖掘问题显著和地球系统模拟预测精度不高等。未来地球大数据的重点研究方向主要集中在：跨学科的地球大数据共享与融合；构建地球科学大数据深度挖掘理论与方法；建设地球大数据智慧平台；形成新的研究范式与协同研究模式。

参考文献

1 Guo H D. Big Earth Data: A new frontier in Earth and information sciences. Big Earth Data, 2017, 1(1/2): 4-20.

2 郭华东. 科学大数据——国家大数据战略的基石. 中国科学院院刊, 2018, 33(8): 768-773.

3 Guo H D, Wang L Z, Liang D. Big Earth Data: A new frontier in Earth and Information sciences. Science Bulletin, 2016, 61(7): 505-513.

4 王淑强, 王卷乐, 李扬, 等. 基于文献计量学的国际地球科学数据管理研究进展. 全球变化数据学报, 2020, 4(3): 299-313.

5 董少春, 齐浩, 胡欢. 地球科学大数据的现状与发展. 科学技术与工程, 2019, 19(20): 1-11.

Big Earth Data Research Topic Evolution and Influence Analysis

AN Peijun[1,3*] XIAO Xiantao[1] GUO Huadong[2] YAN Dongmei[2] LI Jialei[1,3]

（1 Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou 730000, China;

2 Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China;

3 Department of Library, Information and Archives Management, School of Economics and Management,

University of Chinese Academy of Sciences, Beijing 100049, China ）

Abstract Big data driven data-intensive scientific discovery, and the Big Earth Data is the base and engine of Earth system science. Based on observation of natural geographical elements on different scales and acquisition of data, and the data obtained during the interdisciplinary research processes, integration and mining analysis will promote the development of Earth science to achieve significant innovation. The Big Earth Data has also become the important national strategic resources, and exerted influence on international cooperation and diplomatic relationship. In this study by means of literature metrology, the Big Earth Data literature from the Science Citation Index (SCI) and Social Science Citation Index (SSCI) was analyzed. According to the global paper output quantity, the influence of countries and institutions, research topic, research hotspot change, international cooperation, and so on, present situation of the Big Earth Data research was revealed. Some suggestions in the future will focus on sharing and integration of developing interdisciplinary Big Earth Data, and in order to improve the Earth science data depth mining theory and method, and to realize analysis, modeling and prediction of complex Earth system, and to support and serve global change and sustainable development.

Keywords Big Earth Data, literature metrology, distribution of research topics, research hotspots evolution, research influence

*Corresponding author



安培浚 中国科学院西北生态环境资源研究院文献情报中心副研究员。中国科学院青年创新促进会会员，中国科学院“西部之光”人才项目入选者。主要研究方向：科技战略与科技政策研究、地球科学战略情报研究。先后主持国家自然科学基金青年基金、国家自然科学基金国际合作项目、中国科学院战略情报研究能力提升项目和中国科学院政策研究课题等20多项研究项目，发表论文30多篇。E-mail: anpj@llas.ac.cn

AN Peijun Associate Professor at the Information Center, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences (CAS), Member of Youth Innovation Promotion Association of CAS, as well as Winner of the Western Light Talents Training Program of CAS. She has taken charge of several projects including the project sponsored by Young Scientists Fund of National Natural Science Foundation of China, the project of international cooperation, and has published over 30 papers. Her major research directions include science & technology strategy and policy research, strategic intelligence research on earth science. E-mail: anpj@llas.ac.cn

■责任编辑：文彦杰